

eJournal Archiving: Status and Trends

Tyler O. Walters

Associate Director, Technology & Resource Services
Georgia Tech Library & Information Center

ER&L – Feb. 22, 2007



Status of eJournal Management

- One characteristic present today: Anxiety
 - re: preservation / long-term access to EJs
- Arrangements via license agreements are inadequate – how will libraries get content?
- Many EJs not covered by archiving arrangements
- Need technical tools to enforce archiving arrangements in licenses

The Shift to ePublishing

- Ulrich's International Periodicals Directory (June 2006)... tells us there are:
- 23,187 active and refereed journal titles (low)
- 14,338 active, online, and refereed journal titles (62%) (38% print only)
- 1,429 open access titles (10% of online journals)

More on Shift to ePublishing

- Directory of Open Access Journals:
 - 2,549 peer-reviewed OA journals (2/1/07), up 1,100 titles fr. 2005
- British Library study (2003):
 - By 2016 half of all serials will be e-only. Only 9 yrs. from now!
- Karen Hunter, Elsevier (2006):
 - For publishers to move to electronic-only format, we need “bullet-proof digital archiving of eJournals”
- USE: Ithaka study (2003):
 - 78% of 7,400 faculty surveyed: EJs are “invaluable research tools”
- Carol Tenopir (2003):
 - 2/3 of scientist's reading from e-resources; astronomers 80%

What Are Librarians Doing?

- Publishers Communication Group (2004):
 - 155 academic librarians surveyed worldwide, 84% cancel print when e-version is available
- Elsevier Science Direct (2006):
 - 40% of current subscription revenues from e-only subscriptions
- ARL Member Survey (2005):
 - 98% of contracts have provision for some form of backfile access if a library cancels its subscription
But can you trust the publisher to deliver?

Core Infrastructure

Two e-Resource Coordinators

For Collection Development

SELECT

BUDGET

**USE
UTILIZATION**

PROMOTE

EVALUATE

For Acquisition Services

LICENSE

ORGANIZE

ACCESS

**MAINTAIN
TROUBLESHOOT**

PRESERVE

eJournal Archiving Programs

Characteristics to Consider:

1. Mission and Mandate

- Explicitly expressed - legal deposit, legal mandates

2. Rights and Responsibilities

- Right to Preserve: clearly enumerated, remain viable over time
- Are publishers, libraries involved in governance? operations?

3. Content Coverage

- Which publications? for whom? title, issue, date

4. Minimal Service (*Minimum Criteria for an Archival Repository of Dig. Scholarly Journals, DLF*)

- Receive, store, integrity check, processing, threat protection, audit

eJournal Archiving Programs

More Characteristics to Consider:

5. Access Rights

- Current Access vs. Archiving / Dark Archive vs. Light Archive
- Trigger events for access or transfer

6. Organizational Viability

- org. context: responsible admin., sources of funding,
- stakeholder buy-in

7. Network

- “Repositories work as part of network (org & tech). Redundancy
- Share strategies/planning docs, coordinate content selection, reciprocal archiving/mirroring, 2nd ary archiving responsibility

Perpetual Rights Access vs. EJ Archiving

The Programs...

1. CISTI Csi (Canada) (pra)
- 2-3 LOCKSS Alliance and CLOCKSS (eja)
4. KB e-Depot (Netherlands) (eja/pra)
5. Kopal/DDB (Germany) (eja/pra)
6. LANL-RL (eja)
7. NLA PANDORA (Australia) (eja/pra)
8. OCLC ECO (pra)
9. OhioLINK EJC (pra)
10. Ontario Scholars Portal (pra)
11. Portico (eja)
12. PubMed Central (pra)

Recommendations for Libraries

- Press publishers hard to enter into archiving relationships with bona fide programs. Act together to get archiving programs that meet their needs. Put this requirement in your license agreements!
- Share information with each other:
 - what they're doing about e-journal archiving
 - internal assessment processes for decision making,
 - codify best practices
 - promote sufficient redundancy
 - shares responsibility for preserving EJs not currently included
- Join at least one EJ archiving initiative
- Need a registry of archived scholarly publications that indicates which programs have archived them (like ROAR for repositories)

Recommendations for Publishers

- Be overt about EJ archiving efforts, enter into relationships with at least one EJ archiving program
- Provide enough information to EJ archiving programs to ensure that the scope, content, date span, and title coverage are adequately recorded
- Extend liberal archiving rights in license agreements with content aggregators, consortia, libraries. Digital archiving rights of EJs = a distributed responsibility
- **Recommendations = Transparency of Operations**

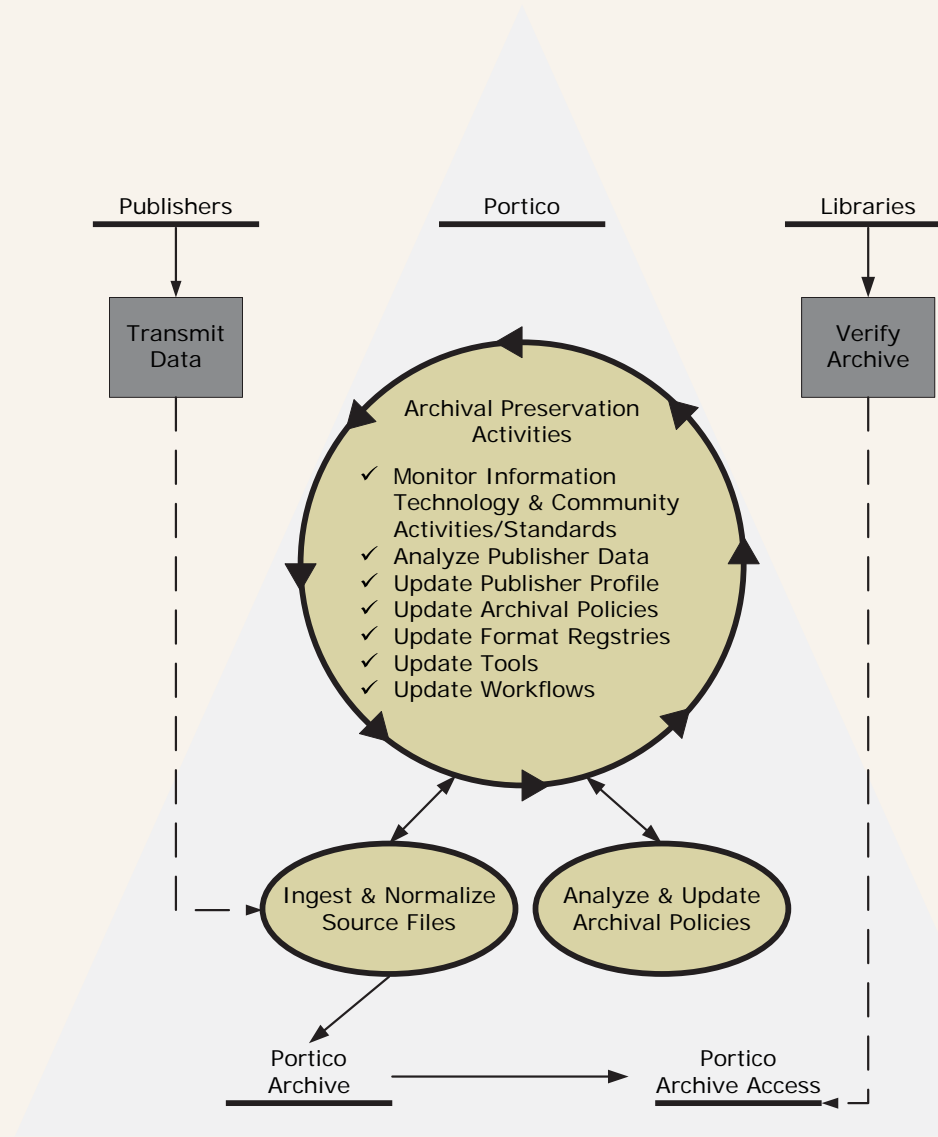
Recommendations for EJ Archiving Programs

- Public evidence of minimal level of services (be transparent)
- Be overt about content in programs, put this info. on WWW
- Content becomes repository property. Can't remove, modify. Disputes
- Include rights / responsibilities in contracts to protect archiving actions
- Consider, some content will enter public domain. Should negotiate all agreements with publishers to take this possibility into account
- Exchange information: content coverage, technical implementations, best practices for rights to preserve and eventually provide access to content. Create a safety net for one another for succession planning
- Share responsibility for preserving EJs not currently protected

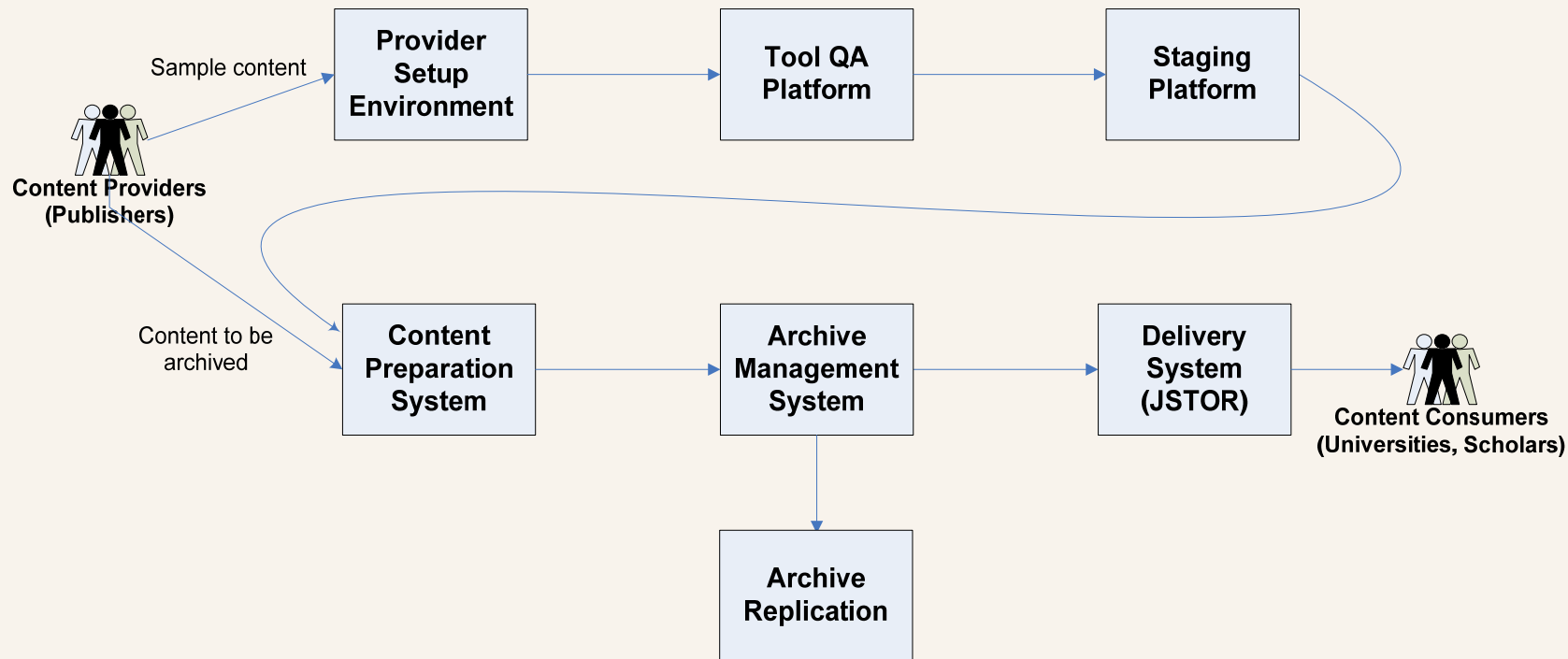
Examples of EJ Archiving Programs:

- Portico
- LOCKSS / CLOCKSS

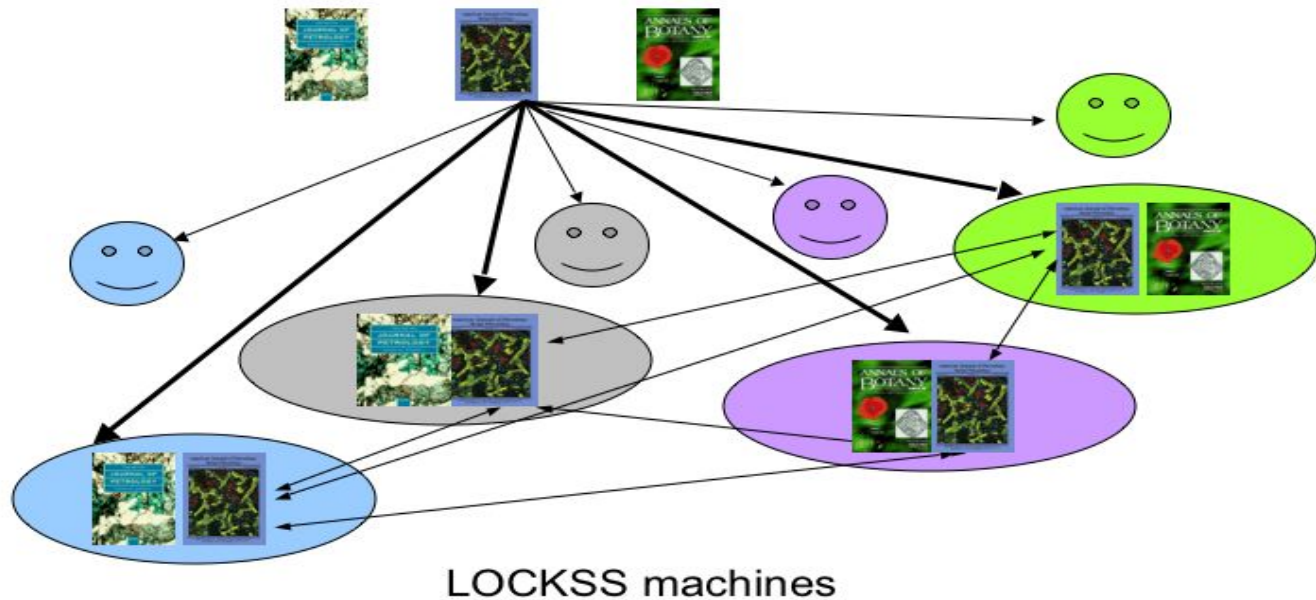
Portico Preservation Process



Ingest Process Overview



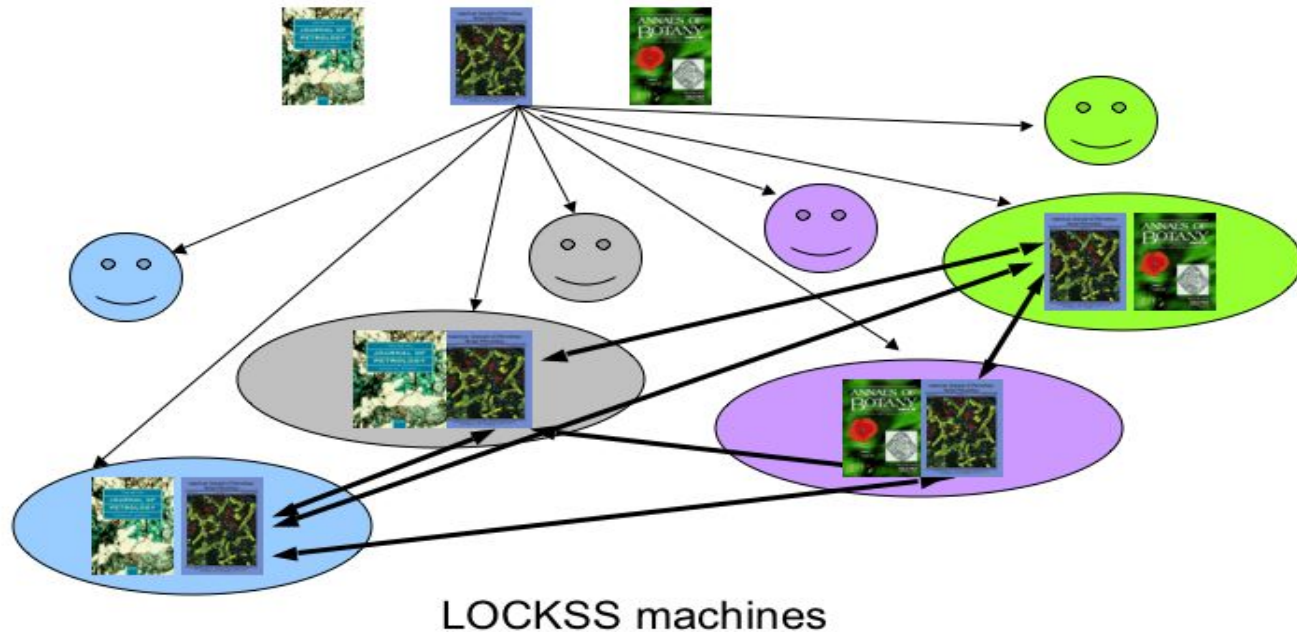
Collection via Web Crawler



Preserving and Auditing

- Uses Internet to continually and slowly audit preserved content
- Servers take part in polls, voting on part of content they have in common. If server content is damaged or incomplete, server will lose poll, and can repair the content from other servers
- Avoids need to back servers up individually. Provides reassurance that system is performing its function and that correct content is available to readers when they access it
- More organizations preserving content = stronger guarantee for their continued access

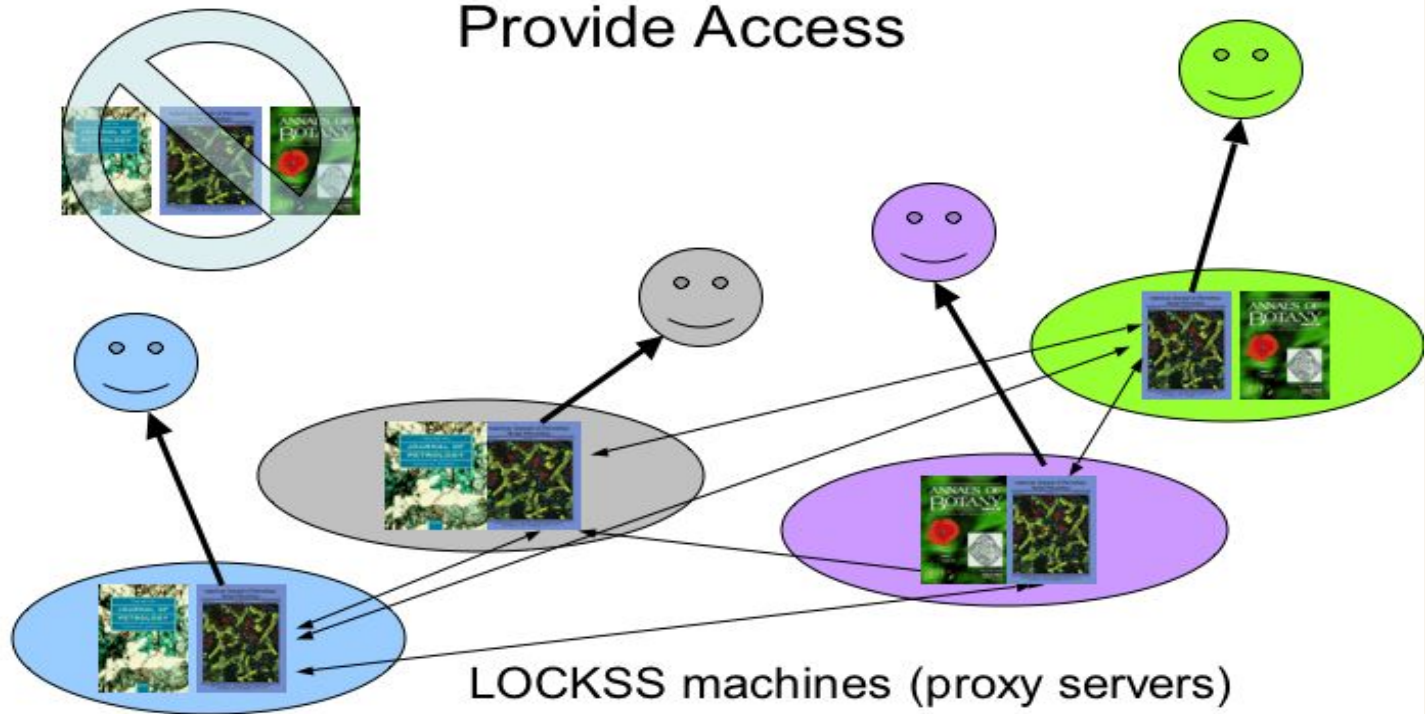
Preserve and audit content integrity



Providing Access

- LOCKSS servers provide transparent access to the content they preserve
- Institutions run web proxies to allow off-campus users to access their journal subscriptions, and web caches, to reduce the bandwidth cost of providing Web access to community
- Integrates with these systems, intercepts requests from community's browsers. When request for page from preserved journal arrives, it is first forwarded to the publisher
- If publisher returns content, browser gets that. Otherwise browser gets preserved copy

Provide Access



Ok... Discussion time!

- *“E-Journal Archiving Metes and Bounds: A Survey of the Landscape,”* CLIR. Nov. 2006

Tyler O. Walters,
Tyler@gatech.edu
404-385-4489